Modelling small area at-risk-of-poverty rates for the UK using the World Bank methodology and the EU-SILC

Patricia C. Melo[a*], Andrew Copus[a], Mike Coombes[b]
[a] Social, Economic and Geographical Sciences, The James Hutton Institute, Craigiebuckler, Aberdeen AB15 8QH, United Kingdom
[b] Centre for Urban and Regional Development Studies (CURDS), Newcastle University, Newcastle upon Tyne, NE1 7RU, United Kingdom

* Author for correspondence: patricia.melo@hutton.ac.uk, Tel. +44 (0)1224 395316.

Abstract

The at-risk-of-poverty rate is a key monitoring indicator in connection with the European Union's goal to take 20 million people out of the risk of poverty and social exclusion by 2020. Whilst there is comprehensive and up to date coverage at the national level, much less is known about how different regions are performing in this respect. This paper illustrates how the World Bank poverty mapping methodology, combined with the European Union Statistics on Income and Living Conditions (EU-SILC) survey, can be used to compute small area estimates of at-risk-of-poverty rates for the UK, and compares the obtained estimates with existing estimates of relative poverty obtained using national surveys. There is considerable spatial dispersion in at-risk-of-poverty rates. The highest rates tend to be found in large cities, but there are also relatively high rates in some remote rural areas. Furthermore, regional differences in housing costs can act as an important driver of poverty, particularly in large cities. Our analysis suggests that the EU-SILC survey, combined with national population census data, can provide a practical basis for developing regional or local estimates of at-risk-of-poverty rates across EU Member States, which would be particularly valuable where national surveys have inadequate regional sampling.

Keywords: small area estimation, at-risk-of poverty rate, poverty mapping, EU-SILC, World Bank

Word count: Total: 6297 words [Abstract: 200 words; Text: 6097 (excludes references)]

1. Introduction

In 2010 the European Union (EU) adopted an overall strategy, known as EU2020, to guide both Community policies and, where the EU has no direct policy competence, the policies of Member States (through the so-called Open Method of Coordination). One of the headline targets of the EU2020 strategy is to lift 20 million people out of poverty by 2020 (EUROPEAN COMMISSION, 2010). Three indicators are to be used in support of this: the number of persons at risk of poverty; the number of persons not able to afford four of the nine items indicative of material deprivation; and the number of persons living in households where all the adults work less than 20% of a full time year. For policy impact monitoring purposes the number of persons in each of these categories are added together (while avoiding double counting of individuals), and each Member State has a separate reduction target which, added together, gives the EU total of 20 million (Eurostat, 2004, 2005, 2007, 2012).

The "at-risk-of-poverty" (ARoP) indicator, which constitutes the first element of the EU2020 target, was adopted by the EU Council as early as 1975. This indicator is defined as the number (or percentage) of people who have a net income of less than 60% of the national median equivalised disposable income (after social transfers). This indicator may be considered rather idiosyncratic when used to make comparisons at a continental scale, due to its dependence upon national benchmarks (Ward, 2009, Bradshaw and Mayhew, 2010). Nevertheless, its wide acceptance renders it a key indicator in a policy context. More immediate than the EU2020 goals, the new programming period (2014-20) for the European Structural and Investment Funds presents some opportunities for regional policy alleviation strategies. Regional targeting generates a demand for more detailed information on regional and local patterns of poverty. Given the current budgetary limitations, it is important that such interventions are carefully targeted on regions where they may have the greatest impact.

At present, Eurostat publishes ARoP rate data for about two-thirds of the countries within the ESPON[1] space at NUTS 2, the remaining countries provide data at NUTS 1 or NUTS 0 (whole

[1] ESPON (European Observation Network, Territorial Development and Cohesion) is the research programme adopted by the European Commission to support place-based policy supported by multiple Member State applied research projects. For more details see the link: http://www.espon.eu/main/ (accessed 7th March 2014).

country).[2] These data are mostly derived from the EU-SILC (Survey of Income and Living Conditions). Sample sizes constrain publication of ARoP rates at a more detailed regional level based upon this source. A few countries, notably the Nordic countries and the Netherlands, are able to generate ARoP rates from administrative registers. In these countries it is possible to generate reliable ARoP rates at NUTS 3 or even smaller areas. Elsewhere, some form of estimation is necessary.

The ESPON 2013 programme's 'TiPSE' project (Territorial Dimensions of Poverty and Social Exclusion in Europe)[3] has been tasked with collating existing regional data on poverty and social exclusion, and with estimation where no secondary data are available. In this paper we draw on some of the work carried out within TiPSE and illustrate how the World Bank (WB) poverty mapping methodology, combined with the European Union Statistics on Income and Living Conditions survey, can be used to estimate model-based at-risk-of-poverty rates for small areas in the UK. The EU-SILC provides comparable microdata on income, social exclusion and living costs across member countries and may therefore provide a way of producing comparable small area estimates of poverty across EU Member States, which is especially valuable where national surveys are not available. We compare our estimates of relative poverty obtained from the EU-SILC survey with existing estimates of relative poverty for the UK obtained from national surveys.

Our findings indicate that the average at-risk-of-poverty rate before and after housing costs is about 15% and 24% respectively, and there is considerable spatial variation in poverty rates. The highest poverty rates tend to be found in large cities, but there are also relatively high rates in some remote and sparsely populated rural areas. Poverty rates based on equivalised disposable income after housing cost are generally higher than poverty rates based on equivalised disposable income before housing cost, particularly in large urban areas. This suggests that regional differences in housing costs can act as an important driver of poverty. Although there

[2] The NUTS nomenclature (abbreviation of the French *Nomenclature des Unités Territoriales Statistiques*) is the geographical classification of the EU territory created by the European Office for Statistics (Eurostat) for the reporting of European level statistics. It divides each Member State into a series of NUTS 1 regions, which in turns are each subdivided into smaller NUTS 2 and NUTS 3 regions (EUROPEAN COMMISSION, 2007).

[3] TiPSE is a three year project, led by Nordregio (Stockholm) with research partners in the UK, Germany, Greece, and Hungary. Further information and documents may be found at http://www.espon.eu/main/Menu_Projects/Menu_AppliedResearch/tipse.html (accessed 10th January 2014).

are some differences between the relative distribution of our estimates of poverty rates and those obtained from national surveys, our analysis suggests that the EU-SILC provides a viable way of generating model-based small area estimates of poverty across EU countries.

The structure of the paper is as follows. Section 2 provides a brief overview of small area estimation, while Section 3 describes the small area estimation methodology used by the World Bank and the data used in the estimation of at-risk-of-poverty rates for the UK. Section 4 presents and discusses the results, and Section 5 summarises the main conclusions.

2. Small area estimation of at-risk-of-poverty rates

Small area estimation (SAE) methodologies have been applied as a way of producing estimates of income and poverty for small geographical areas with limited, or zero, sample size in survey data. The spatial aspect of SAE is particularly relevant here because in the EU policy context poverty is defined by reference to a comparator group, and for SAE models the definition of the comparator population is crucial. In practice the EU requires that national level values are used as the comparator (Guio, 2005), which is notably different to the regional policy context where the whole EU provides the comparator.

Applying SAE techniques to estimate poverty risk at a sub-national scale for which the survey data on income have too small a sample is one way to ‘borrow strength’ from survey responses in nearby areas (Verma et al., 2005). If income survey data are available at the household level – microdata – SAE methods have proved appropriate, with some based on econometric models and others on spatial micro-simulation models. Both methods proceed by using survey microdata (e.g. households, individuals) for a group of variables supposed to be good predictors of income and for which robust data are available from census and/or administrative sources for each of the small areas in the analysis. The main difference between the two methods is the way they link the survey data to census data to estimate small area estimates of income and poverty.

Work based on econometric modelling applies the model parameters obtained from the income regressions based on survey microdata to census micro data using the same set of variables (i.e. income predictors) used in the survey-based regression models. Modelling is generally based on mixed models with area-specific variance components to capture between-area variability. More recently, quantile regression models have been proposed as a way of providing a more detailed representation of income levels by estimating the quantiles of the distribution, instead of only the conditional mean (e.g. Chambers and Tzavidis, 2006, Tzavidis et al., 2008).

Spatial micro-simulation also starts by estimating regression models using survey microdata to identify the variables which correlate well with income, and which are used in a following stage to match survey data to small area census data through a reweighting process. Small area estimates are generated by repeatedly adjusting the survey household weights using small area census data for each of the variables selected as good income predictors, where the weights represent the probability of a given household of living in a given small area (see Ballas et al., 2005 for more details about spatial micro-simulation).

One important limitation of both econometric model- and simulation-based methods is the uncertainty resulting from both survey sampling and model misspecification (Fenton, 2013). There is no consensus over which of the two SAE methods provides the more reliable option, with debate continuing among researchers such as Brinegar and Popick (2010) and, with a specific poverty focus, Molina and Rao (2010). Heady and Hennell (2001) considered alternative approaches for the particular situation posed by poverty across Europe, and their view that a linear regression model is a reasonable 'default' offers some support for the approach here which is based on the WB small area estimation method.

Whereas the two SAE methods use income data to provide an income-based measure of poverty, other approaches to the measurement of poverty use 'proxy' methods because income data are not available. For example the lack of an income question in the UK census has stimulated much research developing multiple-variable indexes of deprivation aiming to capture the different dimensions of poverty and related issues. The most prominent of these is the Index of Multiple Deprivation (Noble et al., 2000, Noble et al., 2006), which was preceded by work such as Carstairs and Morris (1989) and Gordon (1995) and has been followed by Norman (2010) among others. This approach is based on the concept of relative deprivation established in the original work by Townsend (1979). It is an understanding of deprivation as a multifaceted condition that goes beyond income, and is operationalised in broad measures of living standards of which the most know is the *consensual,* or *perceived deprivation,* method (e.g. 1983 Breadline Britain Survey, 1990 Breadline Britain Survey of Britain, 1999 Poverty and Social Exclusion

Survey of Britain – see Mack and Lansley, 1985, Gordon and Pantazis, 1997, Gordon et al., 2000).

One of the main issues for proxy methods is the limited ability to compare indicators across studies and over time as a result of their widely diverging selection of proxy variables. However the analysis of comparable estimates of poverty over time has been achieved in the case of Gordon (1995), Dorling et al. (2007) and Fahmy et al. (2011).

Although the measure of poverty used in this study (i.e. income based at-risk-of-poverty rate) cannot capture the multifaceted expressions of deprivation which extend beyond income and material deprivation, it has the advantage of being calculable for small areas and comparable across numerous countries. This is critically important in the context of the EU2020 strategy for poverty alleviation. Limiting the scope of this analysis to being at risk of poverty, and hence the measurement to the income dimension of deprivation, allows it to focus directly on exploring the potential of using the EU-SILC survey and WB poverty mapping methodology to generate small area estimates of at-risk-of-poverty rates in EU Member States.

The WB developed a model-based SAE methodology (and related software) that enables users to estimate measures of poverty and income inequality for small or medium-sized regions. The methodology is based on regression models of household income with local area effects to account for between area variability (Lanjouw, 2003, Elbers et al., 2003). Survey data covering a sample of small, or medium-sized, geographical units are used to estimate models of the relationship between income and a set of explanatory variables. The estimated model parameters are then combined with a similar set of covariates obtained from census data for the whole population of small geographical units to predict income levels and poverty measures.

Small area estimation models have been used in the UK by the Office for National Statistics (ONS) to produce estimates of average household income for Middle Layer Super Output Areas (MSOA) in England and Wales, based on data from the Family Resources Survey (FRS) (Bond and Campos, 2010). In addition to small area average income estimates, the ONS has also developed model-based estimates of the proportion of households with income below 60% of the national median income for MSOA in England and Wales using data from the FRS and Households Below Average Income (Fry, 2011).

Small area model-based estimation of poverty rates has also been developed for Scotland. Bramley and Lancaster (1998) generated estimates of income for local and small areas in Scotland, while Bramley and Watkins (2013) produced estimates of both income and poverty for local and small areas in Scotland based on data from the FRS, the Scottish Household Survey (SHS), and the survey Understanding Society. The Scottish Government (2010) has also produced estimates of relative poverty from the SHS and FRS for local authorities in Scotland.

Small area estimates of relative poverty for Northern Ireland were produced by Anderson (2009) using spatial microsimulation modelling and data from the FRS. Estimates of relative poverty were generated for each Super Output Area (SOA) and used to compute measures of relative poverty for other more aggregate geographies by the Northern Ireland Statistics and Research Agency (NISRA).

3. Modelling at-risk-of-poverty rates

This section provides a description of the analysis carried out using survey and census microdata and the WB small area estimation methodology to produce estimates of the at-risk-of-poverty rates for the UK. The geographies used correspond to Local Authorities for England, Wales, and Scotland, and Parliamentary Constituencies for Northern Ireland. The measure of relative poverty (i.e. ARoP) is defined as the proportion of households with equivalised disposable income (before and after housing costs) below 60% of the national median value. Equivalised disposable income adjusts disposable income for household size and composition.

3.1. Small area income model

Although many household surveys contain detailed information about household income, they generally provide an insufficient representation of income patterns for small geographical areas due to limited sample size and limited spatial coverage. On the other hand, census data can provide both wide and detailed spatial coverage but lack information about income. To produce small area estimates of relative poverty for the UK we implement the SAE methodology adopted by the WB, developed by Elbers, Lanjouw and Lanjouw (Elbers et al., 2003, Elbers et al., 2002). It combines survey and census data with regression modelling to generate estimates of income and poverty for small geographical areas. Survey data are used to develop a model of household income from which parameters are estimated and applied to comparable census data for which household income data are not available. The predicted income is then used to calculate measures of poverty and/or income inequality for small geographical areas.

The regression model of household income is first estimated using EU-SILC income data (*inc*) and a set of covariates **X** correlated with income and which are available both in the survey and the census. By using only the covariates available in both datasets, the estimated model parameters can be used to generate the mean distribution of household income for any sub-population in the census conditional on the sub-population's observed covariates **X**. The general form of the income model is given by the following equation:

$$inc_{hc} = \mathrm{E}[inc_{hc} \mid X_{hc}] + u_{hc} = \beta X_{hc} + u_{hc} = \beta X_{hc} + \eta_c + e_{ck} \qquad (1)$$

where h denotes the household and c denotes the survey sample region (or cluster) to which the household belongs. u_{ch} is the model error term, which can be decomposed into the terms η_c and e_{ch}, where η_c captures cluster-specific effects and e_{ch} is the remaining error term.

Besides household characteristics, there may be some contextual regional factors which can help explain part of the observed variation in household income (e.g. unemployment rate, ethnic minorities, etc.). When data for such factors are not available to the analyst an appropriate alternative is to include a cluster effect η_c to capture region-specific heterogeneity.

The estimated model parameters are applied to the census covariates to predict household income levels for the whole population of small areas and, combined with bootstrapping techniques, to produce estimates of poverty. For more details on the WB methodology please refer to Elbers et al. (2003).

3.2. Data and variables

The specification of the household income model described above is based on a set of variables that can predict income levels (i.e. income predictors). This approach has been used in previous studies (e.g. Fay and Herriot, 1979, Bramley and Lancaster, 1998, Bramley and Watkins, 2013) and seeks to avoid a 'statistical fishing trip' by selecting variables that represent factors which evidence suggests do influence levels of poverty risk. The purpose of this model is not to identify theoretical causal relationships between the explanatory variables and household income, but solely to generate empirical estimates of household income. To take one particularly strong example, Ward (2009) reports a high level of correlation between income levels and poverty rates. Atkinson et al. (2010) emphasises the strong link between the experience of joblessness and poverty risk at the household scale, and this was followed by the age and gender focus brought into the analysis of Betti et al. (2012). Putting these principles into practice the model developed here sought the following types of variables:

1) Individual / Household demographic characteristics (e.g. household size, family type, age group, marital status).
2) Individual / Household socio-economic characteristics (e.g. education, qualifications, employment status, occupation, car ownership).
3) Housing characteristics (e.g. type of property, property tenure, number of rooms, presence of shared / separate bathroom, presence of central heating).
4) Individual / Household health conditions (e.g. long-standing illness).

The data used in this study come from the 2005 European Union Statistics on Income and Living Conditions (EU SILC) available from Eurostat, and the UK census 2001 Small Area Microdata (SAM) available from the UK data service. We use the 2005 EU-SILC dataset because it is the first year of the survey which has regional identifiers for NUTS 2. The reference population in the two data sources is not the same. The EU-SILC includes all private households and their current members aged 16 and over, and the data available comprises variables both at the household and personal level. The 2001 census SAM data refer to individuals, although some of the variables do effectively reflect household level data (e.g. number of household members, number of rooms, etc.).[4] Hence, the first task was to create a 'population' in the census SAM dataset comparable with that of the EU-SILC survey, by excluding the records of the census that refer to people younger than 16 years old and people who do not live in a private household (e.g. communal establishments).

[4] The ideal Census dataset for this work would be the 2001 Household Controlled Access Microdata Sample (Household CAMS), but it requires a special permission.

In order to identify the predictor variables for the household income model, we examined the list of questions in the EU-SILC survey and in the census SAM dataset. The number of variables common to the two datasets is limited and was further constrained by a number of irreconcilable definitional differences (e.g. classification of qualification levels). The consideration of additional contextual regional variables (e.g. unemployment rates for larger NUTS 3 and NUTS 2 regions) did not improve the goodness of fit of the income model and was therefore abandoned; this is possibly because such aggregate data cannot capture variation across smaller regions.

After identifying the list of potential candidates, we evaluated the compatibility between the survey and census data by testing the similarity of their relative distributions using the chi-squared (χ^2) test with null hypothesis of a similar relative distribution between survey and census variables. Only the variables with a similar enough distribution (i.e. for which we could not reject the null hypothesis) were kept for the subsequent analysis. The full set of tables comparing the relative distribution of survey and census variables and reporting the results for the chi-squared (χ2) of similar relative distributions can be obtained from the authors upon request. The final EU-SILC sample supporting the regression model of household income contained 10,325 records. The composition of the sample is shown in Table 1. The parameters obtained from the EU-SILC income model (discussed in the following section) are then applied to the same census variables to predict income levels for the population of small areas and combined with bootstrapping techniques to produce estimates of poverty. The final census sample to which this procedure was applied contained 2,294,204 records.

[Insert table 1 here]

4. Results and Discussion

In this section we present and discuss the results obtained from the household income model and associated small area estimates of at-risk-of-poverty rates, and compare our estimates with existing SAE estimates of relative poverty for the UK obtained from national surveys.

4.1. Income models and at-risk-of-poverty rates

Table 2 shows the parameter estimates of the income model estimated using the EU-SILC variables described in the previous section. The model explains about 20% and 18% of the variation in household equivalised disposable income before housing cost and after housing cost respectively. Although small, the values of the coefficient of determination are in line with evidence that explanatory power is lower for cross-sectional data models (as is our case), than for pooled and time series data models (Pindyck and Rubinfeld, 1991). This is because the proportion of variance that cannot be explained is generally higher within a group of different individuals than for a single individual observed over a given time period, or a group of individuals observed over a given time period. Overall, the relationship between income and the different covariates is the same in both analyses. The following paragraphs summarise the main results.

Income levels tend to be higher for ages between 30-39 years old, and lower for ages between 20-24, 60-69, and 50-59 years old. Single parent families are associated with lower income levels, while households without children have the highest income. Smaller households appear to be associated with higher income levels. Compared to employed people, both unemployed and inactive people experience lower income. Owning a car is associated with higher income levels.

Households living in smaller properties are associated with higher income. The presence of central heating and a separate bathroom / toilet is associated with higher income, although the effect is not statistically significant for the latter. There is weak evidence of a conclusive relation between property tenure and income; this may result from the classification used in our analysis, which combines rented accommodation at market price and free accommodation in the same category. The results indicate that there is no statistically significant difference in income levels between households who own their accommodation and households who rent their accommodation at a reduced price, while households who rent accommodation at the market price of have their accommodation for free are associated with lower income levels. Finally, having a long-standing illness that limits activities is associated with lower income.

[Insert Table 2 here]

Table 3 provides some basic descriptive statistics of the estimates of at-risk-of-poverty rates. The rates based on income before housing cost are generally lower than those based on income after housing cost, while there is less dispersion in the latter. The median / mean rate is 14% / 15% and 23% / 24%, respectively. This compares relatively well with existing data for 1998/99, which indicates that the proportion of UK households with income below 60% of the median income was 19% when housing costs are not considered (i.e. income before housing cost) and 24% when housing costs are considered (i.e. income after housing cost) (Department for Work and Pensions, 2013). The top ranked 25% of areas have a poverty risk rate at least 65% / 24% higher than he bottom ranked 25% of areas for income before housing cost / after housing cost (i.e. P75/P25). These differences can be observed in Figure 1, which compares the distribution of the ARoP estimates before and after housing costs.

[Insert Table 3 here]

[Insert Figure 1 here]

Figure 2 shows the spatial distribution of the ARoP rates for income before and after housing cost. The values shown in the legend are the quintiles of the distribution of ARoP rates, that is, observations (i.e. areas) are grouped into five groups of equal size. The areas with the highest poverty rates are located in Northern Ireland and Wales (e.g. Central Valleys, Swansea, and West Wales and the Valleys), parts of the North East of England (e.g. Tyneside, Sunderland, and Middlesbrough), parts of the North West of England (e.g. Liverpool, Manchester), parts of the West Midlands (e.g. Birmingham), parts of inner London (e.g. Tower Hamlets, Hackney, Newham), and parts of Scotland (e.g. Glasgow, Dundee, Western Isles). The figure suggests that income poverty is found in large cities although there are also relatively high poverty rates in some remoter and sparsely populated rural areas.

Although accounting for housing cost generally increases ARoP rates, the areas with the greatest increase are located in London (particularly in inner London). The risk of poverty based on income after housing cost is between 15-16 percentage points higher than the risk of poverty based on income before housing cost for the London boroughs of Camden, City of London, Hackney, Hammersmith and Fulham, Haringey, Islington, Kensington and Chelsea, Lambeth, Lewisham, Newham, Southwark, Tower Hamlets, Wandsworth, and Westminster. This geographical pattern conforms to expectations based on the combination of very high housing costs and substantial numbers of people with low incomes in inner London. Conversely, some areas have notably lower relative ARoP rates after housing costs are considered, such as much of eastern Northern Ireland outside the city of Belfast.

[Insert Figure 2 here]

4.2. At-risk-of-poverty rates comparisons

In this section we compare our small area estimates of risk of poverty - based on the WB methodology, national census data and the EU SILC survey - with existing small area estimates obtained from national (i.e. UK) survey data. The purpose of these comparisons is to assess whether in the absence of adequate regional sampling for national household surveys (which is the case for several EU Member States), the combination of the WB poverty map methodology with the EU-SILC survey can potentially provide a viable way of developing small area estimates of at-risk-of-poverty rates across EU Member States.

Although all the measures of relative poverty share the same definition, and can hence be compared, they do not refer to the same time period. The poverty risk rates estimated in this study refer to 2001 (2011 census microdata were not available at the time of the analysis), while existing estimates for England and Wales, Scotland, and Northern Ireland refer to 2007/08, 2005/08 and 2008, and 2003/05, respectively. As a result, it is difficult to make comparisons based on the actual absolute values of relative poverty rates. In addition, there may be differences in the specific SAE method used (e.g. spatial microsimulation for Northern Ireland), which also make comparisons difficult. We therefore focus on the relative distribution of the at-risk-of-poverty rates to assess whether spatial patterns are reasonably comparable across the different measures.

It is not appropriate here to compare these income-based poverty indicators with any of the proxy measures of deprivation (e.g. Norman, 2010). Subsequent research could investigate the factors leading to a divergence in the spatial patterns of, on the one hand, the *risk* of poverty as shown by the ARoP rate here and, on the other hand measures of the deprivation, which is an *outcome* of poverty, along with other factors.

4.2.1 At-risk-of-poverty rates for England and Wales

The estimates of small area poverty rates produced by the ONS for England and Wales are not available at the same geographical level as those estimated in this study (i.e. local authorities). In order to provide a comparison for England and Wales, we calculated comparable estimates of at-risk-of-poverty rates at the level of local authorities by combining the MSOA estimates of relative poverty with counts of households per MSOA from the 2001 census. These estimates are only available for income after housing cost. Estimates for the proportion of households in poverty before housing cost were not released due to greater instability of these estimates (Fry, 2011, p. 2).

Table 4 compares the mean, median, and spread of the after housing cost at-risk-of-poverty rates obtained from our analysis (denoted as WB) with those produced by the ONS (Fry, 2011) for the period 2007/08 (denoted as ONS). The pairwise correlation between the two measures is 0.83.

The mean and median values of the poverty risk rate are 24% and 23% respectively, compared to 20% for the ONS estimates. There appears to be greater dispersion in the distribution of ONS estimates (i.e. higher coefficient of variation). The spread of central values (i.e. P75/P25) indicates that the top ranked 25% of areas have a poverty rate at least 1.35 times higher (or 35% higher) than the bottom ranked 25% of areas for the ONS measure, while the difference is 23% for the WB measure. The spread of extreme values (i.e. P90/P10) indicates that the top ranked 10% of areas have a poverty rate at least 1.80 times higher (or 80% higher) than the bottom ranked 10% of areas for the ONS measure, while the difference is 40% for the WB measure.

[Insert Table 4 here]

Figure 3 shows the spatial distribution of poverty rates using quintiles and reveals some clear similarities between the two measures. On both analyses, the areas with highest risk of poverty are located in the west and south of Wales, parts of the North East, North West, West Midlands, and some of London's boroughs (particularly in inner London). The areas with the lowest risk of poverty are mostly located in the South East and the East of England. Very few areas are given ARoP rates by the two analyses which are so different they are not in the same quintile, or at least one of the 'adjacent' quintiles.

[Insert Figure 3 here]

4.2.2 At-risk-of-poverty rates for Scotland

Table 5 compares the mean, median, and spread of the at-risk-of-poverty rates obtained from our analysis (denoted as WB) with those obtained from Bramley and Watkins (2013) for 2008 (denoted as BW) and from the Scottish Government (2010) for the period 2005/08 (denoted as SG).

Similarly to our analysis, the estimates obtained by Bramley and Watkins (2013) and the Scottish Government (2010) also identify high poverty rates both in urban areas (particularly larger cities, e.g. Glasgow, Dundee) and in remoter rural areas (e.g. Western Isles). The pairwise correlation between our estimates and those produced by Bramley and Watkins and the Scottish Government is 0.76 and 0.57 respectively for income before housing cost, and 0.89 for income after housing cost (estimates available for Bramley and Watkins only).

The mean value of poverty risk rates before housing cost is about 19% for BW and SG, and 17% for WB, while the median value is 18% for BW, 20% for SG and 16% for WB. The coefficient of variation indicates that there is greater dispersion in the WB estimates of poverty, followed by SG and BW. The spread of central values indicates that the top ranked 25% of areas have a poverty rate at least 36% higher than the bottom ranked 25% of areas for the WB measure, while the difference is 17% and 27% for BW and SG respectively. The spread of extreme values indicates that the top ranked 10% of areas have a poverty rate at least 70% higher than the bottom ranked 10% of areas for the WB measure, while the difference is 39% and 44% for BW and SG respectively.

There is no poverty risk measure for equivalised disposable income after housing cost for SG, so we compare only the WB and the BW measures. The mean and median values are both about 19% for the BW measure and about 25% for the WB measure. The spread of central values is very similar between the two measures, 1.15 and 1.16 for BW and WB respectively. The spread of extreme values indicates that the lowest poverty rate in the upper 10% group of areas is 27% higher than the highest poverty rate in the lowest 10% group of areas for the WB measure, while the difference is 41% for the BW measure.

[Insert Table 5 here]

Figures 4 and 5 show the spatial distribution of poverty risk rates for income before housing cost and after housing cost, respectively. The values are ranked using quintiles, each containing 20% of the areas in the sample. Some key patterns can be identified. The risk of poverty is highest in remote rural areas like the Western Isles, but also in urbanised areas, particularly Glasgow and its surrounding areas, and Dundee. On the other hand, the areas with the lowest risk of poverty tend to be located in North Eastern Scotland, Perth and Kinross, and the Shetland Islands.

[Insert Figure 4 here]

[Insert Figure 5 here]

4.2.3 At-risk-of-poverty rates for Northern Ireland

Based on the SOA-level estimates of relative poverty, NISRA computed measures of relative poverty for other, more aggregate, geographies as weighted averages of the SOA figures. These estimates are only available for income before housing cost and refer to the period 2003/05. We compare our estimates of relative poverty (denoted as WB) for Parliamentary Constituencies with those computed by NISRA (denoted as NISRA).

Table 6 compares the mean, median, and spread of before housing cost poverty rates. The pairwise correlation between the two measures is 0.95. The mean and median values are considerably different between measures, 31% for the WB measure and 17% for the NISRA measure. In addition, there appears to be greater dispersion in the distribution of the WB estimates. The spread of central values indicates that the top ranked 25% of areas have a poverty rate at least 35% higher than the bottom ranked 25% of areas for the WB measure, while the difference is 17% for the NISRA measure. The spread of extreme values indicates that the top ranked 10% of areas have a poverty rate at least 69% higher than the bottom ranked 10% of areas for the WB measure, while the difference is 38% for the NISRA measure.

[Insert Table 6 here]

Figure 6 shows the spatial distribution of poverty rates using quintiles. Despite the higher values of relative poverty for the WB measure (see Table 6), the figure suggests that the relative distribution of poverty rates across space appears to be fairly similar between the two measures. The areas with lowest risk of poverty are located in the eastern parts of Northern Ireland, with the exception of Belfast. The areas with highest risk of poverty also tend to be the same for both measures, and include parts of Belfast (Belfast North and Belfast West) and Foyle (which includes the city of Derry/Londonderry).

[Insert Figure 6 here]

5. Conclusions

In order to achieve the European Union's goal of reducing the number of people at risk of poverty and social exclusion by 20 million by the year 2020, it is important to be able to monitor poverty and social exclusion at the regional level within EU Member States. At present, Eurostat data for at-risk-of-poverty rates, based on the EU Survey of Income and Living Conditions, refer only to very aggregate regions (NUTS 2, NUTS 1 and NUTS 0).

In this paper, we evaluate the potential of using the EU-SILC survey to generate model-based estimates of at-risk-of-poverty rates for small or medium-sized areas of EU Member States. We apply the WB poverty mapping methodology, in conjunction with EU-SILC and national census data, to estimate at-risk-of-poverty rates for small areas in the UK. The UK provides a good base for investigating the potential of the EU-SILC survey because the ONS, SG and NISRA have recently started producing their own small area estimates of relative poverty for England and Wales, Scotland, and Northern Ireland, respectively. We can therefore compare our estimates of the at-risk-of-poverty rate with those produced by the ONS, SG, and NISRA using national survey data.

The pairwise correlation between our estimates and those obtained from national surveys are generally strong: 0.83 for England and Wales, between 0.57 and 0.89 for Scotland, and 0.95 for Northern Ireland. Making comparisons of the actual absolute values of the at-risk-of-poverty rate across the different measures is difficult because they all refer to different time periods and in some cases use a different SAE method (i.e. Northern Ireland). Nevertheless, it is possible to make comparisons of the relative distribution of poverty rates to evaluate whether they reveal reasonably similar spatial patterns. The comparison of ARoP rates before and after housing costs suggests that the latter has a significant role to play in large urban areas. It serves as an important reminder that income variations are only part of the explanation of the geography of poverty.

Our main conclusion is that the combination of the WB poverty mapping methodology with EU-SILC data and national census data can provide a practical basis for developing small or medium-sized area estimates of at-risk-of-poverty rates across EU Member States. This is likely to be of special importance for Member States where household income data collected through national surveys are not available or are too limited in sample size at a regional level to be used. In addition, such geographically more detailed at-risk-of-poverty rates will also be important in the context of the new programming period (2014-20) for the European Structural and Investment Funds, particularly in the design of regional policy strategies for poverty alleviation.

Acknowledgements

This research was conducted as part of two research programmes funded by the ESPON European Observation Network for Territorial Development and Cohesion programme and the Scottish Government Rural and Environment Science and Analytical Services Division (RESAS). We would also like to thank two anonymous reviewers for their comments.

References

ANDERSON, B. 2009. Creating small area income deprivation estimates for Northern Ireland: spatial microsimulation modelling. *Centre for Research in Economic Sociology and Innovation (CRESI) Working Paper 2009-07.* Colchester: University of Essex.

ATKINSON, A., MARLIER, E., MONTAINGNE, F. & REINSTADTLER, A. 2010. Income poverty and income inequality. *In:* ATKINSON, A. & MARLIER, E. (eds.) *Income and living conditions in Europe.* Luxembourg: Eurostat.

BALLAS, D., ROSSITER, D., THOMAS, B., CLARKE, G. & DORLING, D. 2005. Geography matters: Simulating the local impacts of national social policies. Joseph Rowntree Foundation.

BETTI, G., GAGLIARDI, F., LEMMI, A. & VERMA, V. 2012. Subnational indicators of poverty and deprivation in Europe: methodology and applications. *Cambridge Journal of Regions, Economy and Society,* 5**,** 129-147.

BOND, S. & CAMPOS, C. 2010. Understanding income at small area level. *Regional Trends,* 42**,** 80-94.

BRADSHAW, J. & MAYHEW, E. 2010. The Measurement of Extreme Poverty in the European Union. Brussels: Directorate-General for Employment, Social Affairs and Inclusion.

BRAMLEY, G. & LANCASTER, S. 1998. Modelling local and small area income distributions in Scotland. *Environment & Planning C: Government and Policy,* 16**,** 681-706.

BRAMLEY, G. & WATKINS, D. 2013. Local Incomes and Poverty in Scotland: Developing Local and Small Area Estimates and Exploring Patterns of Income Distribution, Poverty and Deprivation. Research Report for the Improvement Service on behalf of four Local Authorities (Edinburgh, Falkirk, Fife and Highland) and the Scottish Government.

BRINEGAR, S. & POPICK, S. 2010. A comparative analysis of small area population estimation methods. *Cartography and Geographic Information Science,* 37**,** 273-284.

CARSTAIRS, V. & MORRIS, R. 1989. Deprivation: explaining differences in mortality between Scotland, England and Wales. *British Medical Journal,* 299**,** 886–889.

CHAMBERS, R. & TZAVIDIS, N. 2006. M-quantile Models for Small Area Estimation. *Biometrika,* 93**,** 255-68.

DEPARTMENT FOR WORK AND PENSIONS 2013. Households Below Average Income: An analysis of the income distribution 1994/95 – 2011/12. London, 13 June 2013

DORLING, D., RIGBY, J., WHEELER, B., BALLAS, D., THOMAS, B., FAHMY, E., GORDON, D. & LUPTON, R. 2007. Poverty, wealth and place in Britain, 1968 to 2005. Joseph Rowntree Foundation

ELBERS, C., LANJOUW, J. O. & LANJOUW, P. 2002. Micro-level estimation of welfare. *Policy Research Working Paper Series 2911.* The World Bank.

ELBERS, C., LANJOUW, J. O. & LANJOUW, P. 2003. Micro-Level Estimation of Poverty and Inequality. *Econometrica,* 71**,** 355-64.

EUROPEAN COMMISSION 2007. Regions in the European Union. Nomenclature of territorial units for statistics NUTS 2006 /EU-27 *Eurostat: Methodologies and working papers.* Luxembourg: Office for Official Publications of the European Communities, 2007.

EUROPEAN COMMISSION 2010. Europe 2020: A strategy for smart, sustainable and inclusive growth, Brussels, 3 March 2010, COM(20190).

EUROSTAT 2004. Poverty and Social Exclusion in the EU, Statistics in Focus 16. Luxembourg: Eurostat.

EUROSTAT 2005. Income Poverty and Social Exclusion in the EU25, Statistics in Focus 13. Luxembourg: Eurostat.

EUROSTAT 2007. Measuring progress towards a more sustainable Europe. Luxembourg: Eurostat.

EUROSTAT 2012. 23 % of EU citizens were at risk of poverty or social exclusion in 2010, Statistics in Focus 9. Luxembourg: Eurostat.

FAHMY, E., GORDON, D., DORLING, D., RIGBY, J. & WHEELER, B. 2011. Poverty and place in Britain, 1968-99. *Environment and Planning A,* 43**,** 594-617.

FAY, R. E. & HERRIOT, R. A. 1979. Estimation of income from small places: An application of James-Stein procedures to census data. *Journal of the American Statistical Association,* 74**,** 269-277.

FENTON, A. 2013. Small-area measures of income poverty. *Working Paper No CASE/173.* Centre for Analysis of Social Exclusion (CASE), London School of Economics.

FRY, R. 2011. Understanding household income poverty at small area level. *Regional Trends,* 43**,** 87–102.

GORDON, D. 1995. Census based deprivation indices: their weighting and validation. *Journal of Epidemiology and Community Health,* 49**,** 39–44.

GORDON, D., ADELMAN, A., ASHWORTH, K., BRADSHAW, J., LEVITAS, R., MIDDLETON, S., PANTAZIS, C., PATSIOS, D., PAYNE, S., TOWNSEND, P. & WILLIAMS, J. 2000. Poverty and Social Exclusion in Britain. York: Joseph Rowntree Foundation.

GORDON, D. & PANTAZIS, C. 1997. Measuring poverty: Breadline Britain in the 1990s *In:* GORDON, D. & PANTAZIS, C. (eds.) *Breadline Britain in the 1990s.* Aldershot: Ashgate.

GUIO, A.-C. 2005. Income poverty and social exclusion in the EU25. *Statistics in Focus 13/2005.* Luxembourg: Eurostat.

HEADY, P. & HENNELL, S. 2001. Enhancing small area estimation techniques to meet European needs. *Statistics in Transition,* 5**,** 195-203.

LANJOUW, P. 2003. Estimating Geographically Disaggregated Welfare Levels and Changes. *In:* SILVA, L. P. D. & BOURGUIGNON, F. (eds.) *The Impact of Economic Policies on Poverty and Income Distribution: Evaluation Techniques and Tools.* New York: Oxford University Press.

MACK, J. & LANSLEY, S. 1985. *Poor Britain,* London, George Allen & Unwin.

MOLINA, I. & RAO, J. 2010. Small area estimation of poverty indicators. *Canadian Journal of Statistics,* 38**,** 369-385.

NOBLE, M., PENHALE, B., SMITH, G., WRIGHT, G., DIBBEN, C., OWEN, T. & AL., E. 2000. Measuring multiple deprivation at the small area level: The indices of deprivation 2000. London: DETR.

NOBLE, M., WRIGHT, G., SMITH, G. & DIBBEN, C. 2006. Measuring multiple deprivation at the small-area level. *Environment & Planning A,* 38**,** 168–185.

NORMAN, P. 2010. Identifying change over time in small area socio-economic deprivation. *Applied Spatial Analysis and Policy,* 3**,** 107-138.

PINDYCK, R. S. & RUBINFELD, D. L. 1991. *Econometric models and economic forecasts,* New York, McGraw-Hill.

SCOTTISH GOVERNMENT 2010. Relative poverty across Scottish Local Authorities. Scottish Government. http://www.scotland.gov.uk/Publications/2010/08/26155956/0.

TOWNSEND, P. 1979. *Poverty in the United Kingdom,* London, Allen Lane and Penguin Books.

TZAVIDIS, N., SALVATI, N., PRATESI, M. & CHAMBERS, R. 2008. M-Quantile Models with Application to Poverty Mapping. *Statistical Methods and Applications*, 17**,** 393-411.

VERMA, V., BETTI, G., LEMMI, A., MULAS, A., NATILLI, M., NERI, L. & SALVATI, N. 2005. Regional indicators to reflect social exclusion and poverty. *Final Report VT/2003/45*. Brussels: European Commission.

WARD, T. 2009. The risk of poverty and income distribution at the regional level. *In:* WARD, T., LELKES, O., SUTHERLAND, H. & TÓTH, I. (eds.) *European Inequalities: Social Inclusion and Income Distribution in the European Union.* Budapest: TÁRKI Social Research Institute Inc.

Table 1: Variables included in the household income model using EU-SILC data

Variable	% of sample
Age group	
16-19 years old	0.4
20-24 years old	3.5
25-29 years old	6.5
30-39 years old	18.3
40-49 years old	19.5
50-59 years old	17.8
60-64 years old	7.3
65-74 years old	13.1
75 or more years old	13.5
Household size	
0-1	31.9
2-4	62.4
5+	5.7
Marital status	
never married	23.9
married / re-married	46.0
divorced, separated, widowed	30.1
Family type	
single parent	6.8
married / cohabiting - no children	36.8
married / cohabiting - children	21.7
other household type	34.6
Employment status	
Working	56.9
Unemployed	2.3
Inactive	40.8
House tenure	
Owner	68.3
renter - at market price or free accommodation	17.1
renter - at reduced price	14.6
Car ownership	
Yes	75.4
No	24.6
Property type	
detached, semi-detached or terraced house	80.7
Flat	19.3
Number of rooms	
1-2	0.8
3-4	32.5
5+	66.7
Presence of bathroom/toilet	
Yes	94.1
No	5.9
Presence of central heating / way to keep warm	
Yes	94.1
No	5.9
Presence of activity limiting long-standing illness	
yes	26.2
No	73.8

Table 2: Income model results

Variables	Income before housing cost		Income after housing cost	
Constant	9.4717	***	9.4613	***
Age group (reference: >=75 years old)				
16-19	-0.1857		-0.1728	
20-24	-0.2986	***	-0.3278	***
25-29	0.0165		-0.0241	
30-39	0.0575	*	0.0241	
40-49	-0.0249		-0.0556	
50-59	-0.0539	*	-0.075	**
60-69	-0.1310	***	-0.1673	***
65-74	-0.0237		-0.0193	
Household size (reference: 5+ people)				
0-1	0.1807	***	0.1129	*
2-4	0.1550	***	0.1541	***
Marital status (reference: divorced, separated, widowed)				
never married	-0.0043		0.0055	
married / re-married	-0.0384		-0.0372	
Family type (reference: single parent)				
married / cohabiting - no children	0.3880	***	0.3931	***
married / cohabiting - children	0.1954	***	0.2072	***
other household type	0.1780	***	0.2193	***
Employment status (reference: working)				
unemployed	-0.8592	***	-0.9706	***
inactive	-0.4991	***	-0.5209	***
House tenure (reference: renter- at reduced price)				
owner	-0.0172		-0.0045	
renter - at market price or free accommodation	-0.2797	***	-0.3026	***
Car ownership (yes vs. no)	0.2125	***	0.1991	***
Property type - detached, semi-detached or terraced house (vs. flat)	-0.1013	***	-0.1182	***
Number rooms in house (reference: 5+ rooms)				
1-2	-0.4302	***	-0.5238	***
3-4	-0.0732	***	-0.0891	***
Presence of separate bathroom and/or toilet (yes vs. no)	0.0611		0.0915	
Presence of central heating / way to keep warm (yes vs. no)	0.1420	***	0.1427	***
Presence of activity limiting long-standing illness (yes vs. no)	-0.0451	**	-0.0315	
Observations	10,325		10,325	
F-statistic (test of model's overall significance)	100.36***		86.030***	
R^2	0.202		0.178	
Adjusted R^2	0.200		0.176	

Legend: ***, **, * indicate significance at 1%, 5%, and 10% respectively.

Table 3: Descriptive statistics of ARoP rate estimates

	Min	Percentile 25	Median	Mean	Percentile 75	Max
Before housing cost	5.60	10.94	14.04	15.11	18.10	44.98
After housing cost	17.82	21.25	23.25	24.07	26.34	37.92

Table 4: Comparison of after housing cost ARoP rates, England and Wales

	Mean	Median	CV	P75/P25	P90/P10
ONS ARoP	20.43	20.00	0.23	1.35	1.80
WB ARoP	23.79	23.00	0.15	1.23	1.40

Legend: CV: coefficient of variation, P10: 10th percentile, P25: 25th percentile, P75: 75th percentile, P90: 90th percentile.

Table 5: Comparison of ARoP rates, Scotland

	Mean	Median	CV	P75/P25	P90/P10
Before housing cost					
BW ARoP	18.50	18.48	0.13	1.17	1.39
SG ARoP	19.13	20.00	0.15	1.27	1.44
WB ARoP	16.68	16.22	0.22	1.36	1.70
After housing cost					
BW ARoP	19.35	18.96	0.15	1.15	1.41
SG ARoP	-	-	-	-	-
WB ARoP	25.10	24.78	0.12	1.16	1.27

Legend: CV: coefficient of variation, P10: 10th percentile, P25: 25th percentile, P75: 75th percentile, P90: 90th percentile.

Table 6: Comparison of before housing cost ARoP rates, Northern Ireland

	Mean	Median	CV	P75/P25	P90/P10
NISRA ARoP	17.13	16.95	0.12	1.17	1.38
WB ARoP	30.89	30.55	0.19	1.35	1.69

Legend: CV: coefficient of variation, P10: 10th percentile, P25: 25th percentile, P75: 75th percentile, P90: 90th percentile.

List of Figures:

Figure 1: Distribution of ARoP rates before and after housing costs

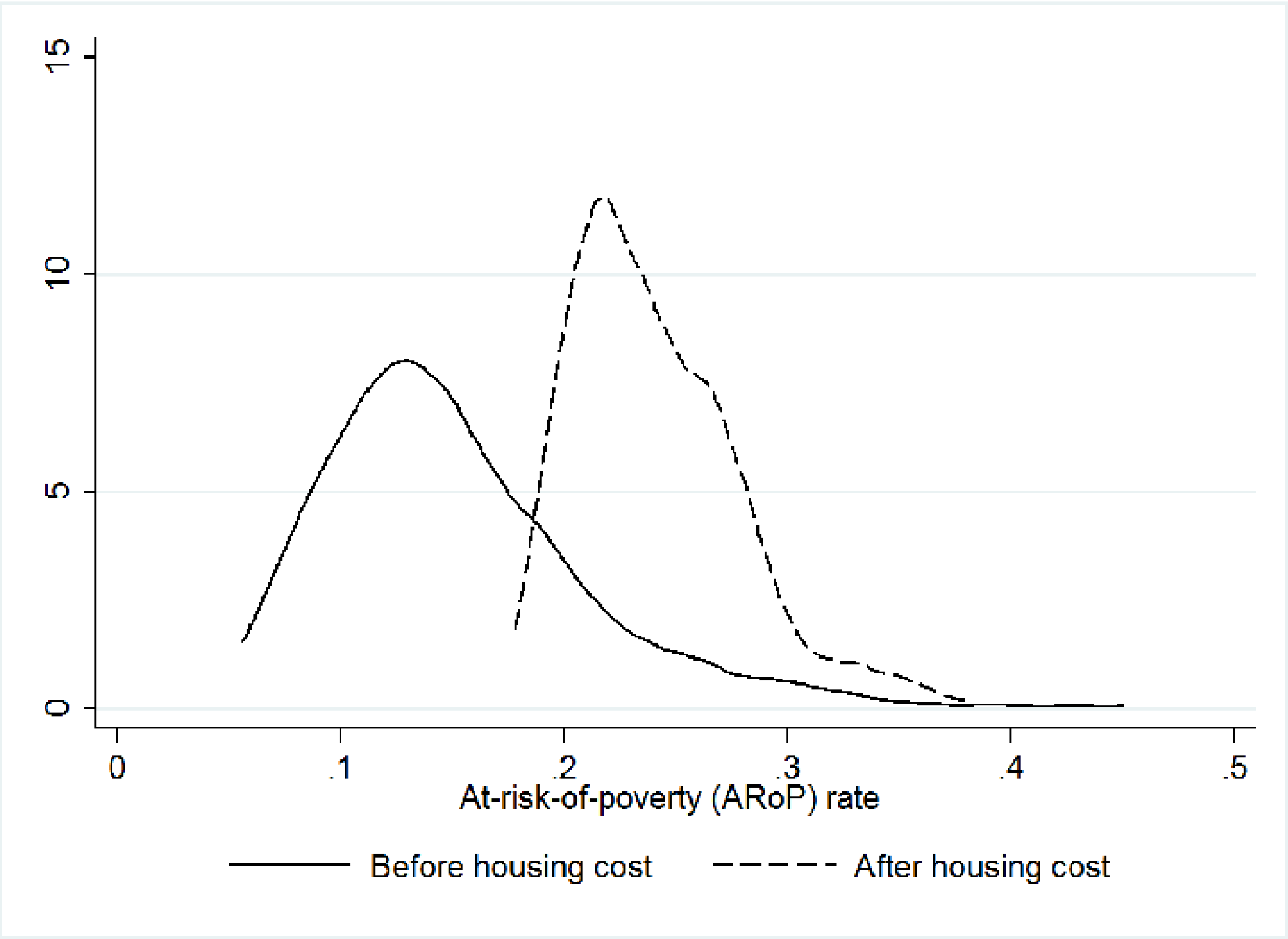

Figure 2: Spatial distribution (quintiles) of ARoP rates before and after housing cost

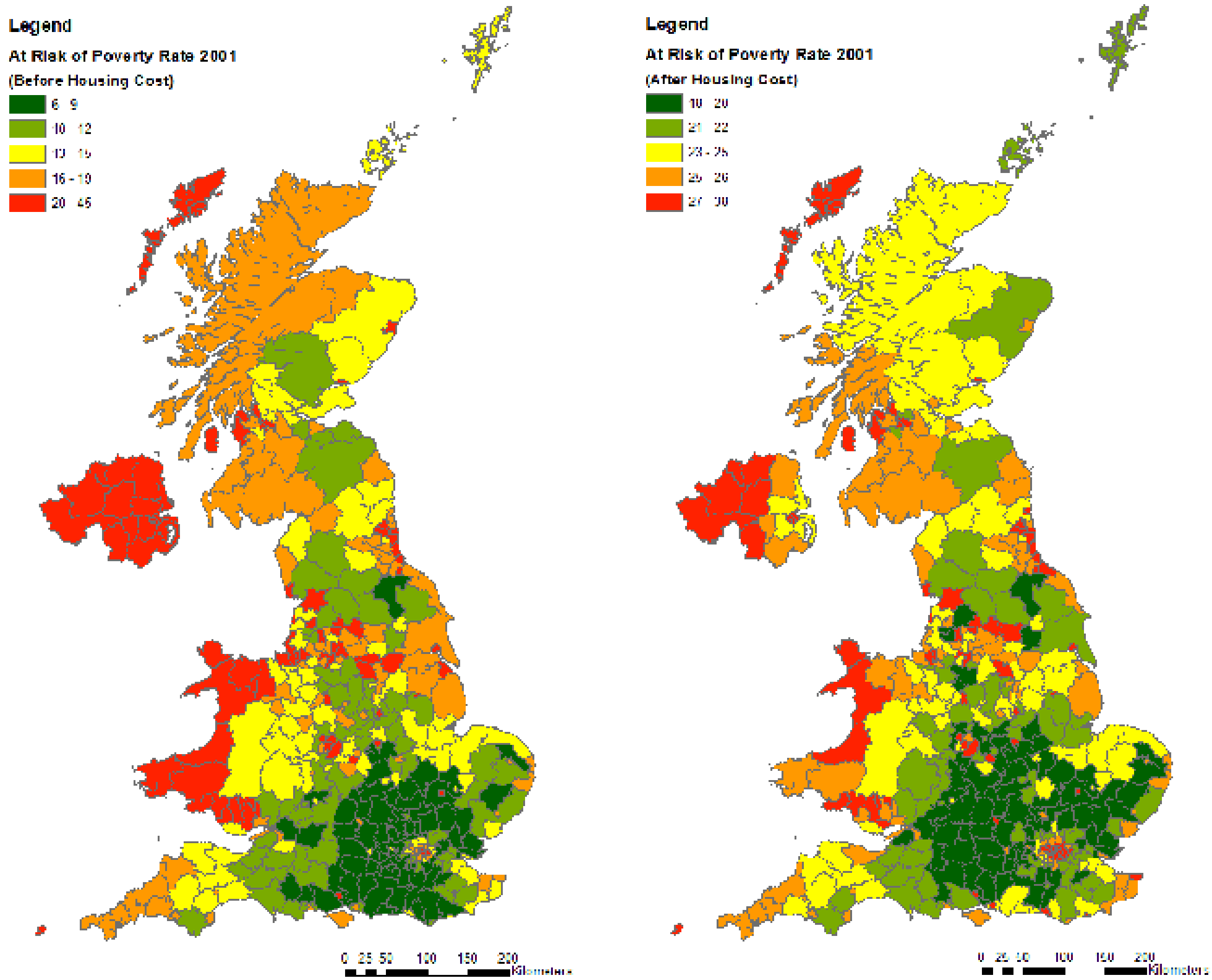

Figure 3: Spatial distribution of ARoP rates after housing cost, England and Wales

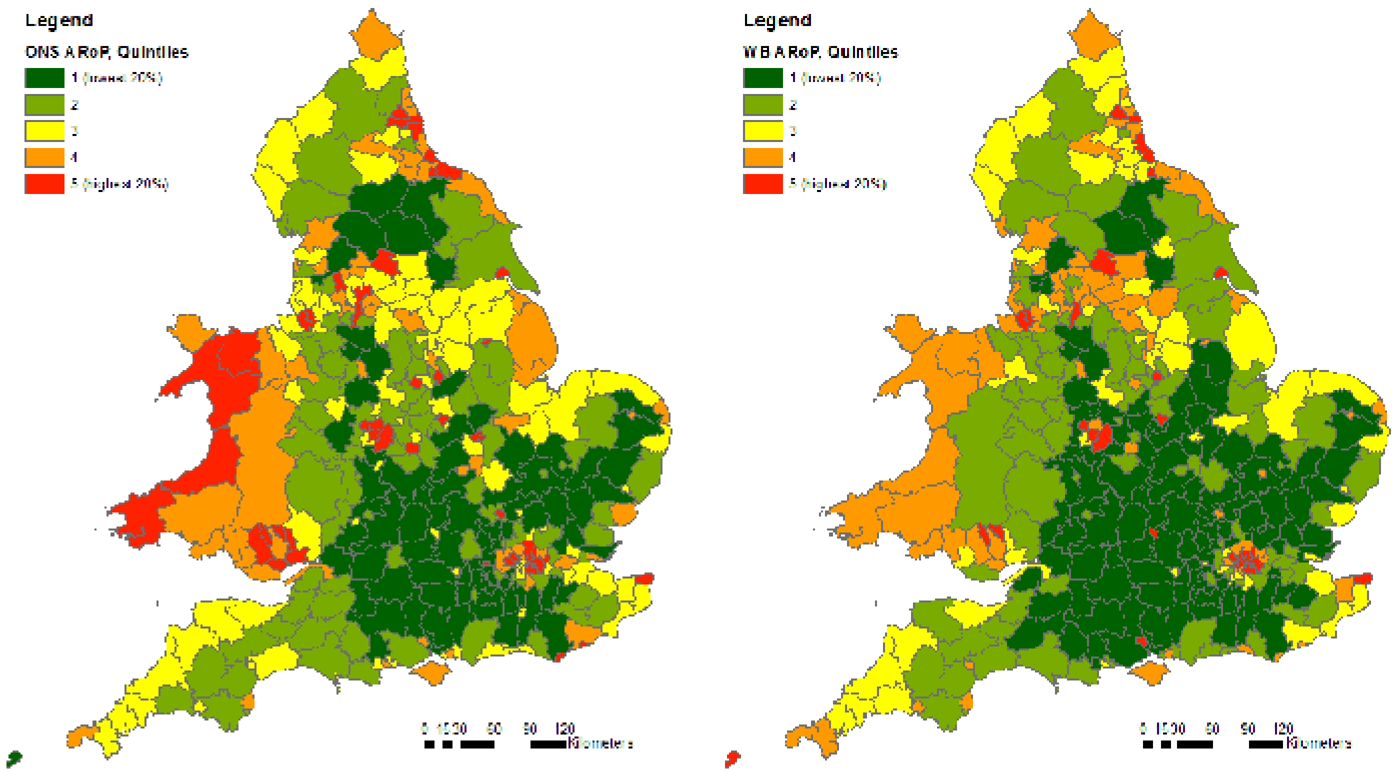

Figure 4: Spatial distribution of ARoP rates before housing cost, Scotland

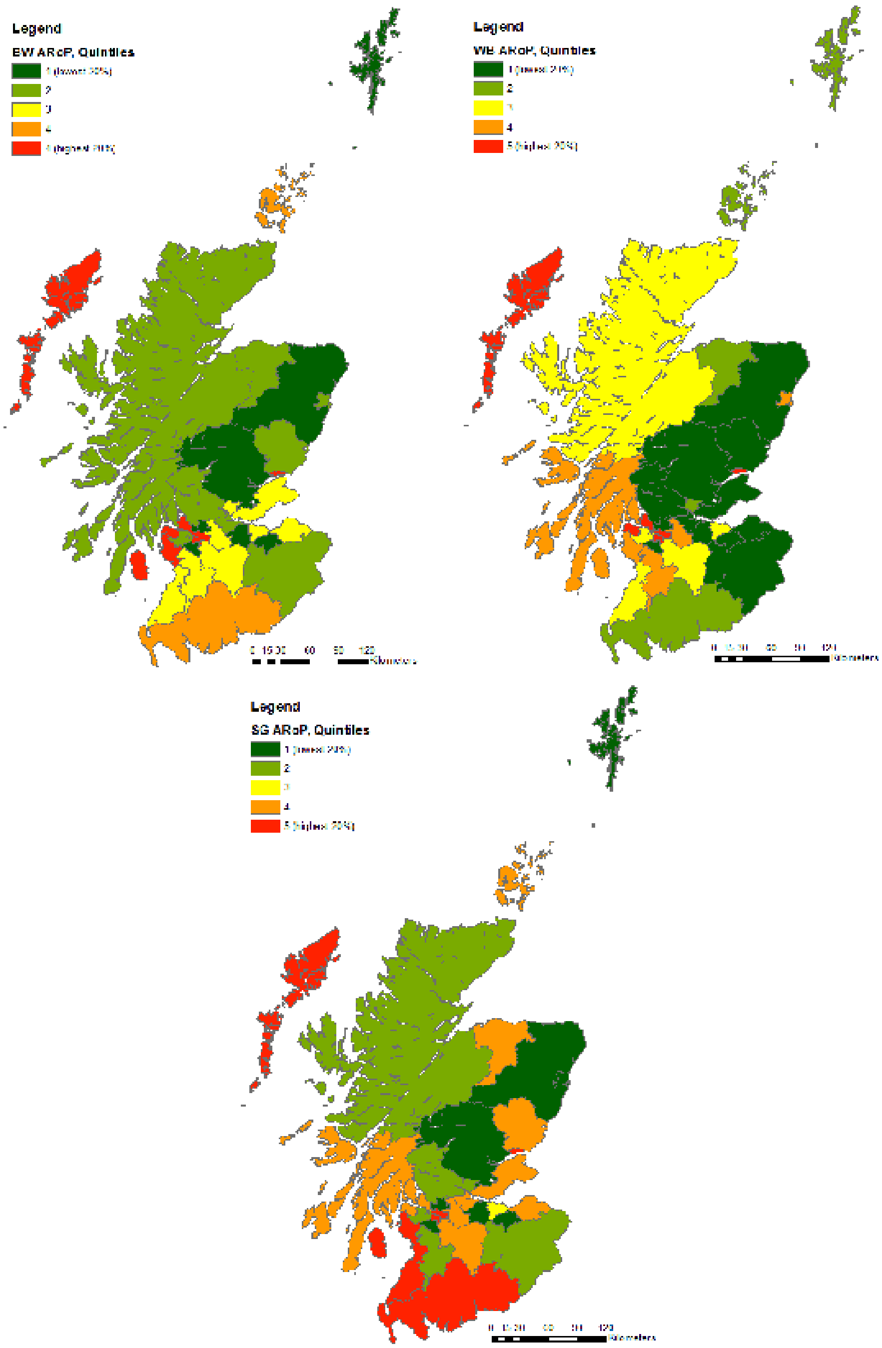

Figure 5: Spatial distribution of ARoP rates after housing cost, Scotland

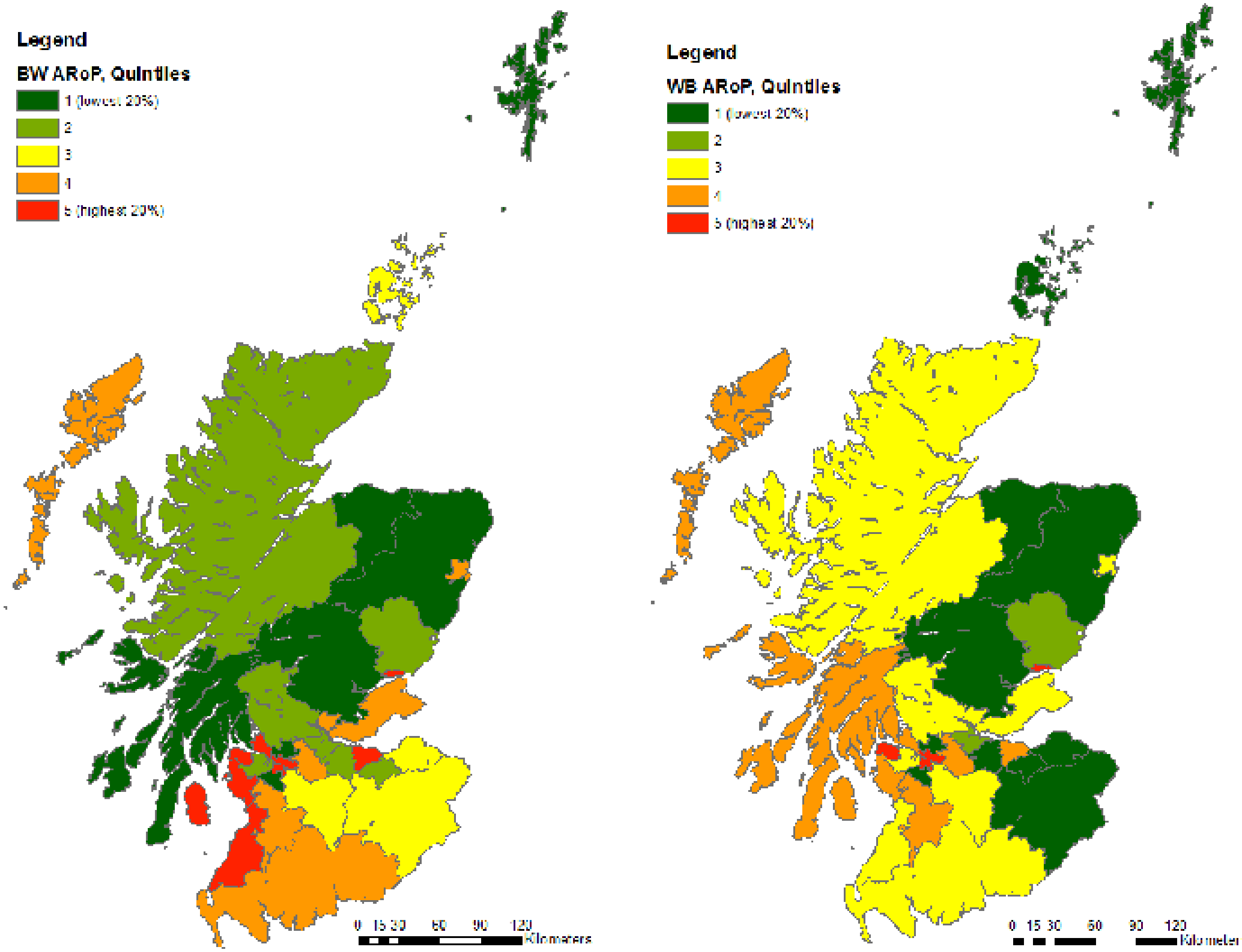

Figure 6: Spatial distribution of ARoP rates before housing cost, Northern Ireland